-43-

Abstract

A method and apparatus is disclosed for reading and decoding information extracted from a form. In the system of the present invention, packages are randomly placed on a conveyor belt, with their labels facing a two-camera subassembly. As the conveyor belt moves, the two-camera subassembly continuously takes images of the belt underneath the overhead camera. The design of the camera permits it to take a high resolution image of a non-singulated, unjustified package flow. A digital image of the packages within the field of view of the camera is then transferred to the processing system for analysis. The processing system identifies individual packages in the image, extracts them and then analyzes the information written on the package labels. The analysis process utilizes conventional Optical Character Recognition (OCR) and Intelligent Character Recognition (ICR) techniques to evaluate the information written on the package label. Once the information is decoded, the system either accesses a database record associated with the decoded machine-readable code, or creates a new record. When an unknown word image is encountered, the field-specific recognition process is aided by use of lexicon information, optimized based on installation-specific or user-specific criteria. The lexicon information is continuously revised based on processed form information. In a preferred embodiment, verified destination addresses associated with a user are alphabetized or rank-ordered based on frequency of occurrence. It is only after the system determines that the originating user is not stored in the database does it resort to the ZIP+4 or similar database to verify a destination address.